On the Possibilities of the Limited Precision Weights Neural Networks in Classification Problems

Sorin Draghici, Ishwar K. Sethi

Vision and Neural Networks Laboratory

Department of Computer Science, Wayne State University,

431 State Hall, Detroit, 48202 MI, USA

Abstract - Limited precision neural networks are better suited for hardware implementations. Several researchers have proposed various algorithms which are able to train neural networks with limited precision weights. Also it has been suggested that the limits introduced by the limited precision weights can be compensated by an increased number of layers. This paper shows that, from a theoretical point of view, neural networks with integer weights in the range [-p,p] can solve classification problems for which the minimum euclidian distance in-between two patterns from opposite classes is 1/p. This result can be used in an information theory context to calculate a bound on the number of bits necessary for solving a problem. It is shown that the number of bits is limited by m*n*log(2pD) where m is the number of patterns, n is the dimensionality of the space, p is the weight range and D is the radius of a sphere including all patterns.

Keywords - neural networks, entropy, classification problems, integer weights, number of bits.

1.Introduction

If neural networks are to be used widely, they have to be suited to hardware implementation which is by far the most cost effective solution for large scale use. One problem related to this foreseeable transition towards hardware is that the actual software simulations of neural networks use floating point arithmetic and either double or simple precision weights. Storing so many bits for each weight and implementing floating point operations would make any hardware implementation unreasonably expensive. Limited precision weight neural networks are better suited for such purposes. This is because a limited precision requires fewer bits for storing the weights and also simpler computations. In turn, this determines a decrease in size of the VLSI chip and therefore, a lower cost for the same performance or, alternatively, a better performance for the same price.

In these conditions, investigating the possibilities of limited precision weights becomes very important. One line of research is to find algorithms able to generate neural networks which use limited precision weights while still being able to solve difficult problems. Various papers try to approach this problem from different angles. [Hohfeld, 1992; Xie, 1991, Coggins 1994; Tang, 1993] use a dynamic rescaling of the weights and a corresponding adaptation of the gain of the activation function. [Hohfeld, 1991a, 1991b; Vincent 1992] rely on probabilistic rounding whereas

[Dundar, 1995; Khan 1994, Kwan 1992, 1993; Marchesi 1990, 1993; Simard 1994; Tang 1993] use weight values which are restricted to powers of two. This latter approach is of particular interest because powers of two values and multiplications of such are particularly easy to represent in binary circuitry.

The question is: how far can this approach be used? Given a problem, what sort of precision should we use so that a solution will still exist? [Khan, 1996b] acknowledges the fact that integer weight neural networks (IWNN) with weight values limited at powers of two lack in capabilities in comparison to the real-valued networks but it also suggests that the weaker learning capabilities of such networks can be compensated by an increase in the number of layers. However, Khan et.al's conclusion comes from empirical experiments. Khan suggested that the answer is affirmative. In this case, how many layers should we expect to need? Given a problem, can one say anything quantitative about the network able to solve the problem? In the following, we shall try to find at least some theoretical answers for these questions.

2.Theoretical considerations regarding the possibilities of limited precision neural networks

In [Beiu, 1996], the author gives an elegant proof for some bounds on the number of bits needed in a classification problem in the general case of real values weights[1]. This paper follows the same line of reasoning and tries to establish similar results for the case of neural networks using limited precision weights.

Proposition 1

Using integer weights in the range [-p, p], one can correctly classify any set of patterns for which the minimum distance between two patterns of opposite classes is $d_{min}=1/p$.

Proof

We first consider the 2 dimensional case. Without loss of generality, we consider all patterns are included in the square [-1,1]. If the patterns are spread outside the unit square, the problem can be scaled with a proper modification of d_{min}.

Firstly, we consider the case p=3.

---

[1] Actually, from the very result presented in [Beiu, 1996], it follows that there *is* a limit on the number of bits per weight and therefore, the weights need not be unlimited real values. However, the proof assumes in its first step (an appropriate translation and rotation) that any dividing hyperplane can be placed in *any* position needed. Therefore, for the proof to hold, the weights need to be able to vary continuosly. A posteriori, after the hyperplanes have been placed and the problem solved, the number of bits for *storing the solution* is reduced and respects the calculated bound.

Figure 1 presents the set of hyperplanes which can be implemented with weights in the set {-3, -2, -1, 0, 1, 2, 3}.

Due to the symmetry of the weights, we only need to consider the triangle (0,0), (0.5,0.5), (0,0.5) determined by Ox, y=1/3 and y=-x+1/3. In this triangle, the largest area is the triangle ABC (0,0), (0,1/3), (1/4, 1/12). In this area, the largest distance between two points inside it, is 1/3-ε . Therefore, any set of patterns of two classes included in [-1,1] and which has the minimum distance between two patterns of opposite classes larger than 1/3 can be classified correctly using these boundaries.

We now consider the same problem for any p. We assume the statement holds for p and show it follows for p+1.

Because the set {-p, -(p-1), ... 0, 1, ..., p-1, p} is included in the set {-(p+1), -p, -(p-1), ... 0, 1, ..., p-1, p, p+1}, all the boundaries present in the figure drawn for p will be also present in the figure drawn for p+1. Therefore, each individual region can only be divided subsequently into smaller regions by the newly added lines corresponding to p+1. We need to show that in the p+1 case, the largest distance in any of the regions is 1/(p+1). However, in the p case, there is just one region where there are internal distances larger than 1/(p+1) and this region is the region R^p determined by Ox, y=1/p and y=-x+1/p. Since regions cannot become larger by increasing p, the only region which needs to be considered is region R.

When p becomes p+1, region R^p will be intersected by y=1/(p+1) and y=-x+1/(p+1) which will determine R^{p+1} and other (smaller) regions. Therefore, R^{p+1} will be the region with the largest internal distance in the new situation (for p+1). However, the largest internal distance in R^{p+1} is 1/(p+1)-ε. QED.

We now consider the n-dimensional case. Again, we assume the statement holds for n and we show that it will then hold for n+1. We want to show that there will be no other possible internal distance when we add a new dimension. For reasons similar to those presented in the 2D case, we can concentrate on the volumes V^n. Without loss of generality, we shall use a figure drawn for the 3D case.

We shall concentrate on the transition from n to n+1. When we add a new dimension (the n+1-th), the statement will be true for any n dimensional figure obtained by choosing any n dimensions out of the n+1 available. In Figure 2, the intersection of the volume V^3 with any 2D space (xOy, xOz, yOz) is the triangle A^2 discussed in the 2D case in which the largest distance is 1/p-ε.

Furthermore, the largest internal distance will always be found along the axes because the space is cut by hyperplanes of the form:

$-x_{n+1}+px_n+\ldots+px_1=0$

which go through the origin and 'slice' the space. In Figure 2, the (hyper)plane:

-x+py+pz=0

intersects xOy in OA_{xy} (-x+p*y=0 or y=1/p*x) and xOz in Oa_{xz} (-x+1/p*z=0 or z=1/p*x).

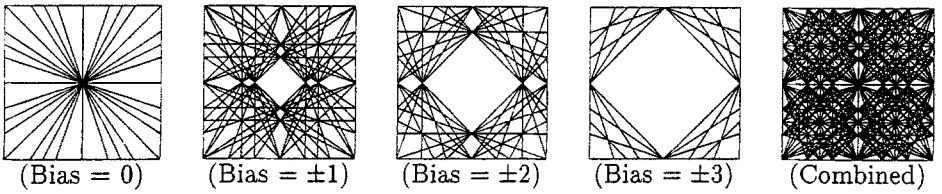

Figure 1 Possible positions for hyperplanes implemented with integer weights in the range [-3, 3] in two dimensions. The picture is drawn for the [-1, 1] square (from [Khan, 1996]).

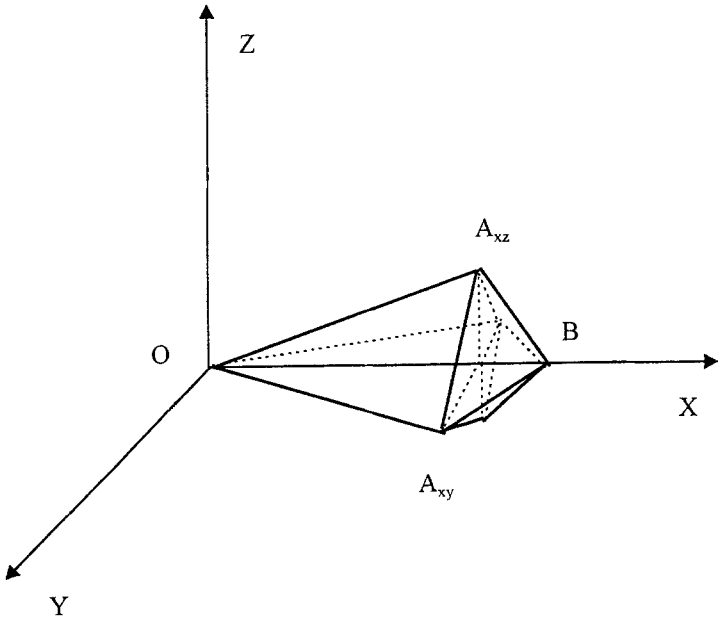

Figure 2 The volume V^3 (in 3D). The largest distance between two points in this volume is OB along the x axis and is $1/p-\varepsilon$.

Because of the hyperplanes similar to $A_{xy}OA_{xz}$ which 'slice' the space radially from the origin, there are 2n volumes V^n each pair of them orientated in the + and - direction along one axis. Each such volume can be imagined as being a union of convex hyperprisms (like $A_{xy}OA_{xz}B$ in Figure 2). All of them have n+2 facets (in n+1 dimensions) and the common OB segment along axis x. But, in each such convex hyperprism (or simplex) the largest distance between two interior points cannot be larger than the longest edge which in this case is precisely OB. Therefore, the largest distance in each such volume V^n will be along the axis associated with that particular volume and it will be still 1/p-ε.

Because the largest distance will remain along the axes, increasing the number of dimensions does not affect the largest distance in one particular volume. Therefore, the largest overall distance is not affected by the added dimension and the largest internal distance in n+1 dimensions will still be 1/p-ε. (However the number of such volumes V^n does increase with the number of dimensions). QED.

3.An entropy bound for the number of bits

As already mentioned, another interesting issue is how complex should the network be for a given problem. Many measures of complexity have been proposed. Among them, there are the depth (the number of edges on the longest input to the output path) and the size (the number of nodes). For VLSI implementation purposes, the depth can be put into correspondence with the delay and the area can be put into correspondence with the area of a VLSI chip. However, these measures are not the best criteria because the area of a neuron depends on the precision of its associated weights. Better criteria are the total number of connections [Hammerstrom, 1988; Abu-Mostafa, 1988; Klaggers, 1993; Phatak, 1994; Mason, 1995], the total number-of-bits needed to represent the weights [Bruck, 1990; Williamson, 1991] or the sum of all the weights and thresholds [Beiu, 1994a, 1995a, 1995b]. The total number of bits is discussed further in [Denker, 1988; Beiu, 1994a] etc. Some entropy bounds for the number of bits have been given by Beiu in [Beiu, 1996] for the case of unlimited precision weights. An similar calculus can be done for the limited precision weights.

We shall consider the case of a set of patterns of two classes with the minimum distance between two patterns from opposite classes d_{min}.

Proposition 2

Let us consider a set of m patterns of two classes in the hypersphere of radius D<=1 centred in origin of $\mathbf{R}^n$. Let us consider $d_{min}=1/p$ the minimum distance between two patterns belonging to different classes. Then, the number of bits necessary for the separation of the patterns (in general positions) using weights in the set {-p, -p-1, ...,0, 1, ...,p} is bounded by

$$\#bits > mn\lceil \log 2pD \rceil$$

Proof

From proposition 1, it follows that one can divide the space using hyperplanes implemented with the given limited precision weights such that the maximum internal distance in any one region is less than 1/p. As shown in the proof of proposition 1, there will be a certain number of 'large' volumes in which the maximum internal distance is 1/p-ε and a number of smaller volumes.

One can calculate the number of bits necessary for the representation of one example p_i as

$$\#bits_{p_i} = \left\lceil \log \frac{V_{total}}{V_{iv}} \right\rceil$$

where V_{total} is the total volume of the problem and V_{iv} is the individual volume of the region which encloses (and separates) the pattern p_i. But all individual volumes are smaller or equal to the 'large' volumes which have the maximum internal distance 1/p-ε. In turn, this volume is convex (from construction) and therefore smaller than the volume V_{hs} of the hypersphere of diameter 1/p. Therefore:

$$\#bits_{p_i} = \left\lceil \log \frac{V_{total}}{V_{iv}} \right\rceil > \left\lceil \log \frac{V_{total}}{V_{hs}} \right\rceil$$

$$\#bits_{p_i} > \left\lceil \log \frac{V_{total}}{V_{hs}} \right\rceil = \left\lceil \log \frac{V_{total}}{\frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} d^n} \right\rceil$$

For n even we have:

$$\#bits_{p_i} > \left\lceil \log \frac{V_{total}}{V_{hs}} \right\rceil = \left\lceil \log \frac{V_{total}}{\frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} d^n} \right\rceil = \left\lceil \log \frac{V_{total}}{\frac{\pi^{\frac{n}{2}}}{\left(\frac{n}{2}\right)!} d^n} \right\rceil$$

But, from hypothesis, all patterns are included in the sphere of radius D centred in the origin. Note that 1/2p<D<1. The first inequality comes from the fact that 1/p is the minimum distance in-between two patterns and 2D is the diameter of the hypersphere containing all patterns. The second one is necessary because proposition 1 was proved for the hypercube $[-1,1]^n$. In these conditions, V_{total} is the volume of the sphere of radius D and the bound can be written as:

$$\#bits_{p_i} > \left\lceil \log \frac{V_{total}}{\frac{\pi^{\frac{n}{2}}}{\left(\frac{n}{2}\right)!} d^n} \right\rceil = \left\lceil \log \frac{\frac{\pi^{\frac{n}{2}}}{\left(\frac{n}{2}\right)!} D^n}{\frac{\pi^{\frac{n}{2}}}{\left(\frac{n}{2}\right)!} d^n} \right\rceil = \left\lceil \log \frac{D^n}{d^n} \right\rceil = \left\lceil n \log \frac{D}{d} \right\rceil =$$

$$= \left\lceil n \log \frac{D}{\frac{1}{2p}} \right\rceil = \lceil n \log 2pD \rceil$$

By multiplying with the number of patterns m:

$$\#bits > mn \lceil \log 2pD \rceil$$

A similar expression can be obtained for n odd.

QED.

4.Discussion

The previous bound must not be interpreted as an absolute lower bound. For a particular problem, when the patterns are in particularly favourable positions, more than one pattern from the same class can share the same volume and thus, the number of bits can be further reduced.

A similar result has been proved in [Beiu, 1996] for the general case in which the weights can have unlimited precision[2] (which allows for a uniform quantization of the space in "elementary" hypercubes of hyperdiagonal d). This allows Beiu to prove the following upper bound:

---

[2] See the previous remark regarding the fact that in [Beiu, 1996] the precision of the weights can be reduced a posteriori.

$$\#bits < mn\left\{\left\lceil \log\left(\frac{D}{d}\right)\right\rceil + \frac{5}{2}\right\}$$

where D is the radius of the hypersphere which includes all patterns and d is the minimum distance in-between the closest patterns from opposite classes. However, the case in which the weights are limited is a particular case of the general case and the Beiu bound must hold. Indeed, in this particular case both bounds hold:

$$mn\left\{\left\lceil \log\left(\frac{D}{d}\right)\right\rceil\right\} < \#bits < mn\left\{\left\lceil \log\left(\frac{D}{d}\right)\right\rceil + \frac{5}{2}\right\}$$

[Beiu, 1997] gives even a tighter bounds for the same number of bits. Assuming there are $m=m_+ + m_-$ examples where m_+ and m_- is the number of patterns in each class respectively then:

$$\min(m_+, m_-) \le m/2$$

and a tight upper bound will be:

$$\#bits < \frac{mn}{2}\left\{\left\lceil \log\left(\frac{D}{d}\right)\right\rceil + 2\right\}$$

which is again consistent with our results:

$$\frac{mn}{2}\left\{\left\lceil \log\left(\frac{D}{d}\right)\right\rceil\right\} < \#bits < \frac{mn}{2}\left\{\left\lceil \log\left(\frac{D}{d}\right)\right\rceil + 2\right\}$$

5.Conclusions

These results show from a theoretical point of view that neural networks using limited precision weights are indeed a viable alternative to neural network using real valued weights. However, the results presented here do not lead directly to an algorithm which is able to actually give the network for a specific problem.

6. Acknowledgments

The authors wish to thank Dr. Valeriu Beiu from Los Alamos National Labs for useful comments on early versions of this manuscript.

7.Bibliography

[Abu-Mostafa, 1988] - Abu-Mostafa Y.S., Connectivity versus Entropy, NIPS'87, D.Z. Anderson (Ed.), Amer. Inst. Of Phys., New York, 1988, 1-8

[Beiu, 1994] - Beiu V., Peperstraete J.A., Vandewalle J., Lauwereins R., Area-time performances of some neural computations, Int. Symp. On Signal Proc., Robotics and NN's, P. Borne, T. Fukuda and S.G. Tzafestas (Eds.), GERF EC, Lille, 1994, pp. 664-668

[Beiu, 1997a] - Beiu V. - Neural Networks Using Threshold Gates: A Complexity Analysis of Their Area and Time Efficient VLSI Representations, Ph.D. thesis, Katholieke Universiteit Leuven, 1994. Extended version to appear as "VLSI Complexity of Discrete Neural Networks", Gordon & Breach, 1997 (in press).

[Beiu, 1995] - Beiu, V., Optimal VLSI Implementations of Neural Networks, Chap. 18 in J.G. Taylor (ed.): Neural Networks and Their Applications, John Wiley & Sons, Chichester, UK, 255-276, 1996.

[Beiu, 1995a] - Beiu, V., Taylor J.G., VLSI optimal neural network learning algorithm, Artif. NN's and Genetic Algs., D.W. Pearson, N.C. Steele and R.F. Albrecht (Eds.), Springer-Verlag, Vienna, 1995, pp. 61-64

[Beiu, 1995b] - Beiu, V., Taylor J.G., Area efficient constructive learning algorithms, Proc. 10th Intl.Conf. on Control Sys. and Comp. Sci., PU Bucharest, Bucharest, 1995, pp. 293-310

[Beiu, 1996] - Beiu, V., Entropy bounds for classification algorithms, Neural Network World, Vol. 6, No. 4, pp. 497-505, IDG Press, 1996

[Beiu, 1997] - Beiu, V., T. De Pauw, Tight bounds on the size of neural networks for classification problems, submitted for IWANN'97

[Bruck, 1990] - Bruck J., Goodman J.W., On the power of neural networks for solving hard problems, NIPS'87, D.Z. Anderson (Ed.), Amer. Inst. Of Phys., NY, 1988, 137-143 (also in J. of Complexity, 6, 1990, 129-135)

[Coggins, 1994] - Coggins R., M. Jabri, Wattle: A Trainable Gain Analogue VLSI Neural Network, Advances in NIPS 6 (NIPS*93, Denver, CO), Morgan Kaufman, San Mateo, CA, 874-881, 1994

[Denker, 1988] - Denker J.S., Wittner B.S., Network Generality, Training Required and Precision Required, NIPS'88, D.Z. Anderson (Ed.), Amer. Inst of Phys., New York, 1988, 219-222

[Dundar, 1995] - Dundar G., K. Rose, The Effect of Quantization on Multilayer Neural Networks, IEEE Transactions on Neural Networks 6(6), pp. 1446-1451, 1995

[Hammerstrom, 1988] - Hammerstrom D., The connectivity analysis of simple associations - or - How many connections do you need?, NIPS'87, D.Z. Anderson (Ed.), Amer. Inst. Of Phys., New York, 1988, 338-347

[Hohfeld, 1991a] - Hohfeld M., S.E. Fahlman, Learning with limited numerical precision using the Cascade-Correlation Algorithm, Tech.Rep. CMU-CS-91-130, School of Comp. Sci. Carnegie Mellon, May 1991. Also in IEEE Transactions on Neural Networks, NN-3(4), 602-611, 1992

[Hohfeld, 1991b] - Hohfeld M., S.E. Fahlman, Probabilistic rounding in neural networks with limited precision. In U. Ruckert and J.A. Nossek (eds.): Microelectronics for Neural Networks (Proc. MicroNeuro'91 - Munich, Germany), Kyrill & Method Verlag, 1-8, October 1991. Also in Neurocomputing, 4, 291-299, 1992

[Khan, 1994] - Khan A.H., E.L. Hines, Integer weight neural networks, Electronics Letters, 30(15), pp. 1237-1238, 1994

[Khan, 1996] - Khan A.H., R.G. Wilson, Integer weight approximation of continuous-weight multilayer feedforward nets, Proc. IEEE Int. Conf. on Neural Networks, vol. 1, pp. 392-397, Washington DC, June 1996, IEEE Press, New York, NY.

[Klaggers, 1993] - Klaggers H., Soegtrop M., Limited fan-in random wired cascade-correlation learning architecture, MicroNeuro'93, D.J. Myers and A.F.Murray (Eds.), Univ.Ed Tech. Ltd. Edinburgh, 1993, pp. 79-82

[Kwan, 1992] - Kwan H.K., Tang C.Z., Designing Multilayer Feedforward Neural Networks Using Simplified Activation Functions and One-Power-of-Two Weights. Electronic Letters, 28(25), pp. 2343-2344, 1992

[Kwan, 1993] - Kwan H.K., Tang C.Z., Multiplierless Multilayer Feedforward Neural Networks Design Suitable for Continuous Input-Output Mapping, Electronic Letters, 29(14), pp. 1259-1260, 1993

[Mason, 1995] - Mason R.D., Robertson W., Mapping Hierarchical Neural Networks to VLSI Hardware, Neural Networks, vol. 8, 6, 1995, 905-913

[Marchesi, 1990] - Marchesi M., G. Orlandi, F. Piazza, L. Pollonara, A. Uncini, Multilayer Perceptrons with Discrete Weights, Proc. Int. Joint Conf. on Neural Networks IJCNN'90, San Diego, Vol. II, pp. 623-630, June 1990

[Marchesi, 1990] - Marchesi M., G. Orlandi, F. Piazza, A. Uncini, Fast Neural Networks without Multipliers, IEEE Transactions on Neural Networks, NN-4(1), pp. 53-62, 1993

[Phatak, 1994] - Phatak D.S., Koren I., Connectivity and performance tradeoffs in the cascade-correlation learning architecture, IEEE Trans. NN's, 5, 6, 1994, 930-935

[Tang, 1993] - Tang C.Z., H.K. Kwan, Multilayer Feedforward Neural Networks with Single Power-of-Two Weights. IEEE Trans. On Signal Processing, SP-41(8), 2724-2727, 1993

[Vincent, 1992] - Vincent J.M., D.J.Myers, Weight Dithering and Wordlength Selection for Digital Backpropagation Networks, BT Technology J., 10(3), pp. 124-133, 1992

[Williamson, 1991] - Williamson R.C., Entropy and the complexity of feedforward neural networks, NIPS'90, R.P. Lippmann, J.E.Moody and D.S. Touretzky (Eds.), Morgan Kaufmann, San Mateo, 1991, pp. 946-952

[Xie, 1991] - Xie Y., M.A. Jabri, Training Algorithms for Limited Precision Feedforward Neural Networks, SEDAL TR 1991-8-3, School of EE, University of Sydney, Australia, 1991. Also in Proc. of the Australian Conference on Neural Networks, Canberra, Australia, 67-71, 1992